DOCUMENT RESUME

ED 197 510 EC 131 712

AUTHOR Mirkin, Phyllis K.: And Others
TITLE Formative Evaluation: Continued Development of Data Utilization Systems.
INSTITUTION Minnesota Univ., Minneapolis. Inst. for Research on Learning Disabilities.
SPONS AGENCY Bureau of Education for the Handicapped (DHEW/OE), Washington, D.C.
REPORT NO IRLD-RR-23
PUB DATE Jan 80
CONTRACT 300-77-0491
NOTE 96p.: For related documents, see EC 131 709-720.

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS *Evaluation Methods: Exceptional Child Research: *Formative Evaluation: *Learning Disabilities: Resource Room Programs

ABSTRACT

A study was conducted to investigate teacher effectiveness when formative evaluation procedures were implemented in their learning disabilities (LD) resource programs on a regular bases. Procedures included two measurement frequencies (daily or weekly) and two means of determining when to change a student's program (decision rules or teacher judgment). Results indicated that daily measurement was significantly more effective than weekly measurement, and that under certain conditions, decision rules were more effective than teacher judgment. Implications are discussed for the development of useful formative evaluation procedures. (Author)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED197510

University of Minnesota

Research Report No. 23

FORMATIVE EVALUATION: CONTINUED DEVELOPMENT OF DATA UTILIZATION SYSTEMS

Phyllis K. Mirkin, Stanley L. Deno, Gerald Tindal, and Kathryn Kuehnle

IRLD

Institute for Research on Learning Disabilities

EC131712

IRLD

Director: James E. Ysseldyke

Associate Director: Phyllis K. Mirkin

The Institute for Research on Learning Disabilities is supported by a contract (300-77-0491) with the Bureau of Education for the Handicapped, Department of Health, Education, and Welfare, U.S. Office of Education, through Title VI-G of Public Law 91-230. Institute investigators are conducting research on the assessment/decision-making/intervention process as it relates to learning disabled children. Research activities are organized into eight major areas:

I. Adequacy of Norm-Referenced Data for Prediction of Success

II. Computer Simulation Research on the Assessment/Decision-making/Intervention Process

III. Comparative Research on Children Labeled LD and Children Failing Academically but not Labeled LD

IV. Surveys on In-the-Field Assessment, Decision Making, and Intervention

V. Ethological Research on Placement Team Decision Making

VI. Bias Following Assessment

VII. Reliability and Validity of Formative Evaluation Procedures

VIII. Data-Utilization Systems in Instructional Programming

Additional information on these research areas may be obtained by writing to the Editor at the Institute.

The research reported herein was conducted under government sponsorship. Contractors are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official position of the Bureau of Education for the Handicapped.

Research Report No. 23

FORMATIVE EVALUATION: CONTINUED DEVELOPMENT OF DATA UTILIZATION SYSTEMS

Phyllis K. Mirkin, Stanley L. Deno, Gerald Tindal, and Kathryn Kuehnle

Institute for Research on Learning Disabilities

University of Minnesota

January, 1980

4

SEP 2 2 1980

Abstract

A study was conducted to investigate teacher effectiveness when formative evaluation procedures were implemented in their learning disabilities resource programs on a regular basis. Procedures included two measurement frequencies (daily or weekly) and two means of determining when to change a student's program (decision rules or teacher judgment). Results indicated that daily measurement was significantly more effective than weekly measurement, and that under certain conditions, decision rules were more effective than teacher judgment. The implications of these findings for the development of useful formative evaluation procedures are discussed.

Formative Evaluation: Continued Development of

Data Utilization Systems

P.L. 94-142 mandates development of an Individual Education Plan (IEP), for all children identified for special education services, that includes educational objectives as well as annual review and periodic monitoring of progress towards goal achievement (Federal Register, 1977). Since periodic review is appraisal of student programs while they are in progress, the type of evaluation required is what Scriven (1967) has termed "formative" rather than "summative" evaluation. Formative evaluation, which is concerned with the evaluation of instructional programs still in some stage of development, appears to more closely approximate the intent of the periodic review process than does summative evaluation, which is concerned with evaluation of a program upon completion. The contrast between formative and summative evaluation is simply and sharply represented by the two questions, "How are we doing?" and "How did we do?" The former, formative question, requires a current and future perspective, while the latter, summative question, requires an historical viewpoint.

Formative evaluation is intervention oriented (Klein, Fenstermacher, & Aiken, 1971) and assumes that systematic assessment and decision making during intervention can substantially improve program effectiveness and affect our summative decisions. The research reported here is part of a larger research program that has as its purpose the development of formative evaluation systems for teachers to use in improving learning disabilities service programs. The present study examined the effectiveness of several components of a formative evaluation system.

The nature of the relationship between progress assessment and program modification in usual classroom or resource room practice is unknown. Presumably, evaluation is informal, nonsystematic, and based on teacher judgment concerning the general quality of student performance on both daily assignments and class participation. Measurement, when it occurs, is likely to take the form of teacher-developed criterion-referenced tests that are administered at intervals determined by the teacher, by the requirements of periodic review, or by the completion of a unit of study. Such tests usually are administered more for the purpose of evaluating student success than for evaluating instructional program success. Using formally developed tests even for such purposes has serious limitations (Jenkins, Deno, & Mirkin, 1979). The most common results of teacher assessment are grades indicating that the student has passed or failed the course or the unit of study, review or reteaching of the items missed on the test, or adjustments in an IEP objective.

In the last decade, alternatives to common teacher assessment practices have been explored. Haring and Berman (1973) note that "the teacher who relies exclusively on before and after pictures of teaching effectiveness is in the position of perenially closing the barn door after the horse is stolen." Along with a number of other researchers (Lindsley, 1971; Lovitt, 1975; Starlin, 1970; White, 1974) he argues that what the teacher needs is "an ongoing measure of effectiveness that permits him/her to alter unsuccessful methods and materials while the student is still able to profit from the alterations." A growing body of literature supports this view (Bohannon, 1975; Liberty, 1972, 1975; Mirkin & Deno, 1979; White & Liberty, 1974).

What is proposed is the design of a system that (a) provides frequent, if not daily, data on student performance, and (b) ensures that the collected data are utilized to make changes in the program when objectives are not being achieved at a satisfactory rate.

Bohannon (1975) compared teacher judgment as the predominant data utilization procedure with Liberty's (1975) specific rules for determining when to make program changes in a beginning reading and phonics program. He found that when teachers used the decision rule treatment, their students achieved more objectives than when teachers used their own judgments. For some students in the Bohannon study, improvement in performance was obtained using only one minute of daily measurement. For these children the decision rules never required program changes. In a related study, Mirkin and Deno (1979) contrasted (a) daily practice on IEP goals in oral reading, (b) daily practice with daily measurement, (c) daily practice with daily measurement and data utilization based on decision rules, and (4) students' regular reading programs. They obtained results that favored structured data utilization based on decision rules.

While research evidence has begun to develop that supports the importance of a system in which teachers use data, a variety of questions remain unanswered. In the study by Bohannon and that by Mirkin and Deno, actual program changes were determined by the experimenter rather than by the teacher. Further, specific program changes were confounded with differences in approach to formative evaluation. The results did not permit generalizations concerning the separate effects of the program changes used, or of the type of decision system used. Since repeated

measurement of performance and data utilization is time consuming, it is important to ensure that any recommended data utilization system is effective when teachers are implementing the instructional strategies routinely used in the classroom program. In that regard, the present research was designed to answer three questions: First, is it more effective to measure student performance daily rather than weekly? Second, when performance is measured daily, is it more effective to have teachers use their own judgment about when and how to change programs or to have teachers change programs in accord with specific decision rules? Third, are there differential effects for different measurement frequencies or for different data utilization strategies?

Method

Sample

Participants were 22 volunteer special education resource teachers and 80 second through sixth grade students classified as learning disabled. Students were of low and middle income socioeconomic status and resided in Minneapolis or St. Paul. Students were randomly selected from a list representing each teacher's caseload of students currently receiving 15 minutes of daily spelling instruction as part of their Individual Education Plan (IEP). All students were at least two years below age grade placement in spelling achievement. The number of students per teacher ranged from two to seven. Sixty students were male and 20 were female. Nineteen teachers were female, three were male, and all had taught special education students for a minimum of three years prior to the study.

Instruments

Two dependent measures were employed: a Grade Specific Test for

each grade level, and a Grand Master Test across all grade levels (see Appendix A). The Grade Specific Tests were constructed by randomly selecting 100 words from grade level word lists from six basal reading series (Harris & Jacobson, 1972). Two such sets of tests were constructed to assess student performance at the conclusion of each experimental period. Three parallel forms of the Grand Master Test, consisting of 60 words each, were constructed by randomly selecting from the entire corpus of 5,167 words in the Harris and Jacobson list.

Tests were administered by graduate research assistants prior to treatment and after the first and second experimental periods. On each occasion, testing included three 3-minute individual dictations of words from the Grade Specific Test representing the student's instructional placement[1] and one 3-minute individual dictation from the Grand Master Test. The 100-word Grade Specific Test was divided into four columns of 25 words each. The starting place for each test dictation was determined using a list of random numbers. Following the first 3-minute dictation, the examiner dictated words from a second list and so forth. The examiner said the word and allowed the student 15 seconds to write the word before the next word was dictated. If the student finished before 15 seconds had passed, the next word was presented immediately. A minimum of 12 words was dictated in each 3-minute period. The same dictation procedure was used to administer the Grand Master Test.[2]

---

[1] To increase sample size at the end of the first treatment period, students placed at the preprimer or primer level were tested from a combined list of preprimer and primer wrods. At the end of the second treatment period students placed at the primer level were tested from a combined primer-Grade 1 list.

[2] The reliability and validity of testing spelling achievement in this way has been well documented elsewhere (Deno, Mirkin, Lowry, & Kuehnle, 1980).

In addition, to assess spelling performance during each experimental period, 15 tests of 25 words each were developed, for each grade level, for each treatment period. These tests were constructed by random sampling (with replacement) from the 100-word Grade Specific Tests (see Appendix B for examples of daily tests)

Placement

Prior to the first experimental period each student's performance in spelling words was assessed on the Grade Specific Tests and used to make instructional placement decisions. Placements were determined by the rate of letter sequences spelled correctly (see Appendix C for scoring procedures). Letters correct per minute in sequence is recommended by White and Haring (1976) as the appropriate level of analysis of student performance in spelling. Starlin (1973) recommends 15 to 29 letters correct/minute as the appropriate instructional level for remedial students.

Procedures

Each teacher was asked to organize a 15-minute daily spelling period for participating students on the total domain of 100 words from the Grade Specific Test representing the level at which the student had been placed. Instructions to teachers emphasized that they were to make all decisions concerning program variations, such as the number of words introduced each day, the order in which words were introduced, and the instructional strategies used to teach the words.

To encourage the teachers to periodically change the instructional program, a representative list of 12 spelling interventions was distributed to each teacher in a checklist format. Teachers were instructed to check the strategies they used for each student daily, and to list

any other procedures they used as well (see Appendix D).

Three different formative evaluation systems were designed and implemented as treatments. Teachers were trained by graduate research assistants to implement the experimental procedures during a one and one-half hour workshop prior to each treatment period (see Appendix E). The formative evaluation systems were as follows:

Daily measurement and data-based rules (DMDB). In this condition teachers divided the daily 15-minute spelling period into 10 minutes for instruction and five minutes for testing. For each student a weekly spelling goal was established by adding 30 percent to the median pretest score on the Grade Specific list for the first week and 30 percent to each weekly goal for successive weeks. The goal was set in terms of letters spelled correctly in sequence and graphically represented on an equal interval graph by a straight line from week to week similar to procedures described by Liberty (1972, 1975). A sample graph with this daily "aim line" is shown in Figure 1.

Insert Figure 1 about here

Following daily spelling instruction the teacher dictated words from one of the Grade Specific lists for a 3-minute period, and then counted with the student the number of letters written correctly in sequence. At a later time, when the student was not present, the teacher plotted the results of the daily measure on the graph. The teacher was instructed to compare the student's performance to the aimline each day. When on three consecutive days, data points were plotted above the aimline, or on three consecutive days, data points were plotted below the aimline, the teacher drew a new aimline on the graph parallel to the original aimline. If the student's performance fell below the aimline for three consecutive days,

in addition to drawing a new aimline, the teacher was instructed to implement a different teaching strategy. If performance was above the line, no new teaching strategy was implemented.

Daily measurement and teacher judgment (DMTJ). The same measurement procedures were used as in the DMDB treatment. In the DMTJ treatment, however, rules were not specified regarding when to change strategies or implement different procedures. Teachers were required to plot the student's daily spelling performance on a graph but a daily aimline was not drawn. Instead, the teacher was asked to judge whether the student's progress was sufficient to continue using the same teaching methods, or whether a new strategy should be invoked to increase performance.

Weekly measurement and teacher judgment (WMTJ). The measurement of performance occurred only once during the week in this system and the student's scores were recorded in a grade book. As in DMTJ, the teacher was asked to judge whether the student's progress was sufficient to continue using the same teaching methods or whether a new strategy should be invoked to increase performance.

Assignment to Treatments

Teachers were randomly assigned to use one of the formative evaluation treatments for a period of three weeks. For the second three week period half the teachers in each treatment were randomly reassigned to one of the other two treatments. As a result, at the conclusion of the study, each teacher had participated in two of the three treatments. The combinations of treatments were as follows:

1) Daily measurement and data-based rules (DMDB)
   Daily measurement and teacher judgment (DMTJ)

2) Daily measurement and data-based rules (DMDB)
   Weekly measurement and teacher judgment (WMTJ)

3) Daily measurement and teacher judgment (DMTJ)
Weekly measurement and teacher judgment (WMTJ)

Seven teachers were in the first treatment combination and eight teachers were in each of the other two combinations. Each combination represents two treatment conditions (e.g., DMDB-DMTJ and DMTJ-DMDB for the first combination), resulting in a total of six different treatments.

Data Analysis

Data were analyzed using covariance analysis on a set of planned comparisons. The effects of treatments were compared both within and between subjects. Analyses included (a) between group analyses of teacher and student data for the three and six week Grand Master and Grade Specific means, (b) between group analyses of teacher and student data for the six week Grand Master means, and (c) within group analyses of both types of data for three and six week data combined. The dependent measures were: rate of letter sequences spelled correctly, rate of words spelled correctly, percent of words spelled correctly, and percent of letter sequences spelled correctly. Sample mean scores were adjusted on the basis of pretest scores to control for any possible differences in groups resulting from random assignment of teachers but not students to treatments.

Results

Between group analyses were conducted to answer the first two questions posed in this study: (a) Are teachers more effective when they measure student performance weekly or daily? and (b) When performance is measured daily does teacher effectiveness differ when changing programs based on their own judgment as opposed to changing programs

in accord with specific decision rules? Within group analyses were conducted to determine the differential effects of different measurement frequencies and different data utilization strategies within treatment groups.

Frequency of Measurement

The question of the effects of different frequencies of measurement was answered using three between group comparisons. One comparison was between student spelling performance when measured daily versus student performance when measured weekly, during the first three-week treatment period. The second comparison was between daily and weekly measurement for the second three-week treatment period. A third analysis compared students who had been measured daily for both treatment periods with students who had been measured daily during one treatment period and weekly during the other treatment period. Covariance analyses were conducted with the student and the teacher[3] each used as the unit of analysis.

Since covariance analysis assumes that the regression of pretest on posttest or vice versa is equivalent across all treatment groups, slopes were computed for the three treatment groups at the beginning of each treatment period and a test for the equivalence of slopes was conducted. The results indicated that the interaction was not significant for either teacher or student data for either treatment period.

Daily vs. weekly measurement in the first treatment period. Table 1 shows the adjusted and unadjusted means and the standard deviations for rate of correct letter sequences and rate of words spelled correctly on

[3]The mean score for each teacher was derived by adding the scores for all the students in each teacher's class and dividing by the number of students.

the Grade Specific and Grand Master Tests when the student was the unit of analysis.

Insert Table 1 about here

The results of the covariance analyses conducted on these means are shown in Table 2. As can be seen in this table, when the means for rate of correct letter sequences were compared, reliable differences favoring daily measurement were obtained on both the Grade Specific and the Grand Master Tests. When the means for rate of words spelled correctly were compared, reliable differences favoring daily measurement were found on the Grade Specific Test but not on the Grand Master Test. When the teacher was the unit of analysis these results were confirmed for the rate of correct letter sequences ($t = 2.53$, $p = .02$) and rate of words spelled correctly ($t = 2.26$, $p = .04$) on the Grade Specific Test. Reliable differences were not obtained for the teacher data, however, between mean rate of words spelled correctly on the Grand Master Test. The p value for correct letter sequences was .06 (see Appendix F, Table 1).

Insert Table 2 about here

To determine whether the differences between daily and weekly measurement reflected an increase in spelling accuracy as well as an increase in the frequency of letter sequences and words spelled correctly, analyses were conducted on the differences between the mean proportions of correct letter sequences and words spelled correctly for students in daily measurement versus students in weekly measurement. The difference between means for the proportion of words spelled correctly on the Grade Specific Test

was found to be highly reliable and favored daily measurement (t = 2.71, p = .008); the difference between the mean proportions of correct letter sequences was somewhat lower (t = 1.77, p = .08). The differences between the means for the proportion of correct words and letter sequences on the Grand Master Test were not reliable (see Appendix F, Table 2). When the same analyses were conducted for the teacher proportion data, these results were confirmed. Reliable differences were found between means for the proportion of words spelled correctly on the Grade Specific Test (t = 2.43, p = .03). Analysis of the difference between means for the proportion of correct letter sequences yielded a p = .08. The obtained differences for either measure on the Grand Master Test were very unreliable (see Appendix F, Table 3).

In summary, of the eight analyses conducted on the rate or proportion of correct letter sequences and words spelled correctly for students in daily measurement versus students in weekly measurement on the Grade Specific Test during the first treatment period, all clearly favored daily measurement. Of the eight comparisons made using Grand Master Test performance means, only two clearly favored daily measurement. Both of these were for rate of letter sequences spelled correctly.

Daily vs. weekly measurement in the second treatment period. The analyses conducted on the first treatment period data were repeated for the second treatment period data. The adjusted and unadjusted means and standard deviations for rate of letter sequences spelled correctly and rate of words spelled correctly on the Grade Specific and Grand Master Tests when the student was used as the unit of analysis are shown in Table 3.

Insert Table 3 about here

The results of the covariance analyses conducted on these means are shown in Table 4. Inspection of this table reveals reliable differences that favor daily measurement for all four comparisons. When the teacher was used as the unit of analysis these results were confirmed for the letter sequence means (t = 2.87, p = .009) and word means (t = 2.32, p = .03) on the Grade Specific Test and for the word means (t = 2.34, p = .03) on the Grand Master Test, but not for the letter sequence means on the Grand Master (see Appendix F, Table 4).

Insert Table 4 about here

Analyses of the proportion data for the second three-week treatment period, when the student was used as the unit of analysis, revealed that differences for both letter sequence and word means were reliable on the Grade Specific Tests (t = 3.12, p = .002; t = 2.05, p = .05), but not on the Grand Master Test (see Appendix F, Table 5). Analysis of the differences for the teacher proportion data for the second three-week treatment period revealed that only the letter sequence difference on the Grade Specific Test was reliable (t = 2.52, p = .02) (see Appendix F, Table 6).

In summary, the analyses conducted on the differences between the rate or proportion of correct letter sequences and words spelled correctly on the Grade Specific Test for students in daily measurement versus students in weekly measurement for the second treatment period revealed reliable differences that favored daily measurement for seven of the eight comparisons. Two of the four differences on Grand Master Test rate means

were reliable and favored daily measurement and one of the differences between proportion means proved to be reliable.

Six weeks of daily measurement versus three weeks of daily measurement and three weeks of weekly measurement. In this analysis, the mean score on the Grand Master Test for students in daily measurement during both treatment periods (DMDB-DMTJ or DMTJ-DMDB pairs) was compared with the mean score for students in daily measurement during one treatment period and weekly measurement during the other treatment period (DMDB-WMTJ, WMTJ-DMDB, DMTJ-WMTJ, and WMTJ-DMTJ pairs). Table 5 shows the adjusted and unadjusted means and standard deviations for rate of letter sequences and words spelled correctly on the Grand Master Test when the student was used as the unit of analysis.

Insert Table 5 about here

As shown in Table 6, reliable differences were revealed between means for rate of letter sequences spelled correctly, again in favor of daily measurement. Differences between means for rate of words spelled correctly were not reliable. When the teacher was used as the unit of analysis these results were confirmed. The difference between means for rate of letter sequences spelled correctly was reliable ($t = 2.16$, $p = .04$) and favored daily measurement, while the difference between means for rate of words spelled correctly yielded a $p = .17$ (see Appendix F, Table 7).

Insert Table 6 about here

Analyses conducted on the differences between means for the proportion of letter sequences and words spelled correctly (for both the student and teacher data) revealed inadequate reliabilities for all except the proportion of correct letter sequences when the teacher was used as the unit of analysis ($t = 1.83$, $p = .08$) (see Appendix F, Tables 8 and 9).

In summary, when the differences between the means for the rate or proportion of correct letter sequences and words spelled correctly on the Grand Master Test for students in daily measurement for both treatment periods were compared with students in daily measurement for one treatment period and weekly measurement for the other treatment period, two differences between means for rate of correct letter sequences were reliable. One comparison between means for proportion of letter sequences spelled correctly approached reliability, while none of the comparisons between means for rate or proportion of words spelled correctly revealed differences that were reliable.

In all, the results of the analyses on the Grade Specific Test data revealed all eight differences between means for rate of letter sequences and words spelled correctly to be reliable favoring daily measurement. Seven of the eight differences between means for proportion of letter sequences and words spelled correctly were also quite reliable. Seven of 12 differences in means for the Grand Master were reliable for rate of correct letter sequences and words. Again, the differences favored daily measurement. Only one of the differences between proportion means on the Grand Master Test appeared to be reliable.

Data Utilization

The question of the effects of different data utilization strategies was answered using two between group analyses for the first and second treatment period data. The means for teachers who measured daily and used their own judgment to decide when to change students' programs were compared with the means for teachers who measured daily and used data-based rules to decide where to change students' programs. The dependent measures again were rate of letter sequences and words spelled correctly, and proportion of letter sequences and words spelled correctly, on the Grade Specific and Grand Master Tests. The teacher and the student were each used as the unit of analysis. The means and standard deviations used in these analyses are those in Tables 1 and 3.

The results of the covariance analyses using mean rate of letter sequences and words spelled co[illegible]tly for the first and second treatment periods when the student was the unit of analyses are displayed in Table 7. As can be seen, none of the differences between means was reliable. These findings were confirmed when the differences between means were analyzed using the student proportion data (see Appendix F, Table 10), the teacher rate data (see Appendix F, Table 11), or the teacher proportion data (see Appendix F, Table 12).

Insert Table 7 about here

In summary, the analyses conducted on the differences between means for students in daily measurement with data-based rules and students in daily measurement with teacher judgment revealed no reliable differences between groups for any of the measures in either treatment period regardless of the unit of analysis.

Within Group Analyses

To determine whether groups of students were differentially affected by different frequencies of measurement or different data utilization methods, three within group analyses were conducted. For the within group analysis, the scores obtained by all the students for a particular treatment in a treatment pair during either the first or second treatment period were combined and a mean score obtained. The obtained scores for each treatment in the pair were then compared to determine whether there were differences between means that were reliable. Table 8 contains the means and standard deviations for rate of letter sequences and words spelled correctly on the Grade Specific Test when the student was the unit of analysis.

Insert Table 8 about here

Table 9 displays the results of the repeated measures analyses conducted on these means. When the means for rate of correct letter sequences were compared, the differences between daily measurement data-based rules and daily measurement teacher judgment or weekly measurement teacher judgment were reliable and favored daily measurement data-based rules; the difference between daily measurement teacher judgment and weekly measurement teacher judgment were not reliable. When the differences in means for rate of words spelled correctly were compared, only the daily measurement data-based rules versus weekly measurement teacher judgment contrast showed a reliable difference.

Insert Table 9 about here

When the student proportion data for letter sequence means were analyzed, these results were confirmed for the daily measurement data-based rule treatment when compared with weekly measurement teacher judgment (see Appendix F, Table 13). Similar comparisons when the teacher was the unit of analysis revealed no reliable differences (see Appendix F, Tables 14 and 15).

In summary, when the means within each treatment pair were compared, students wrote more correct letter sequences and words per minute and achieved greater accuracy in correct letter sequences when instructed under conditions of daily measurement and data-based rules than when instructed under conditions of weekly measurement with teacher judgment regarding when to make program changes. Students also wrote more correct letter sequences per minute when instructed under daily measurement data-based rules than when taught under conditions of daily measurement and teacher judgment regarding when to make program changes.

Discussion

In the present study two components of a formative evaluation system were investigated: frequency of measurement, and data utilization. The specific questions that were posed were: (a) are teachers more effective when they measure student performance weekly or daily, (b) when performance is measured daily does teacher effectiveness differ when changing programs based on their own judgment as opposed to changing programs in accord with specific decision rules, and (c) are there differential effects for different frequencies of measurement and data utilization strategies?

The results of the present study provided clear evidence that teachers were more effective when they measured student performance daily rather than weekly. The results were not as clear with respect to the specific decision rule-teacher judgment contrast.

The evidence favoring daily measurement was most powerful when the results of the analyses for the Grade Specific rate and proportion means were examined. Thirteen of the 16 comparisons revealed reliable differences and two comparisons approached reliability favoring daily measurement. The consistency of this finding, for both the rate and proportion data, suggests that the effect of daily measurement was to increase accuracy as well as frequency of spelling responses. This finding is of particular interest in light of the often cited concern that daily measurement places undue emphasis on "production" at the expense of "accuracy."

The results of the analysis of the Grand Master sample means revealed the Grand Master list to be less sensitive to treatment differences. While more than half the comparisons of the rate means yielded reliable differences favoring daily measurement, the results were somewhat less clear than they were for the Grade Specific sample means. This was particularly the case when the results of the proportion data were examined. Only one of the 12 comparisons approached reliability. When the brief duration of the treatments and the difficulty level of the Grand Master Test are considered, the fact that any differences were found appears to strengthen the case for daily measurement as an important component of formative evaluation.

The evidence to support teacher utilization of specific decision rules or teacher judgment as determinants of when to change student

programs was considerably more tenuous. Although the means for the daily measurement data-based rule treatment were consistently higher than the means for the daily measurement teacher judgment treatment, these were not reliable for any of the comparisons made between treatment groups. The differential effect of treatments on groups of students in treatment pairs, however, revealed reliable differences favoring daily measurement data-based rules over weekly measurement teacher judgment when student data were compared, but not when teacher data were compared. In addition, one of the eight comparisons made between daily measurement data-based rules and daily measurement teacher judgment also revealed a reliable difference that favored the data-based rule treatment. Thus, while the trends of these data appear to favor the daily measurement data-based rule treatment, sufficient evidence was not provided by this study to categorically rule out or include this component, as presently constituted, in a formative evaluation system.

In the present study, teachers in the teacher judgment treatment were provided only minimal training in data analysis for the purpose of making program decisions, while teachers in the data-based rule group were told only to watch the aimline and the daily data points to determine when a program change was needed. The simplicity and ease of implementation of these procedures made them particularly attractive candidates for inclusion in a formative evaluation system. Previous studies, such as those by Bohannon (1975) and Mirkin and Deno (1979) provide additional support to the data-based rule treatment as an effective component of a formative evaluation system. The present results, while confirming the effects for daily measurement found by Bohannon (1975), are somewhat at variance

with previous research findings by both Bohannon and Mirkin and Deno (1979) with respect to the efficacy of the decision rule as a data utilization strategy. The critical variable in the previous studies may have been the consequences that were tied to the decision rule treatment, rather than the decision rule itself. While further research to test this hypothesis would be desirable, the results of the present study suggest that a more fruitful area of research with respect to the development of an effective formative evaluation system may be to test alternative data utilization approaches that include more intensive teacher training in monitoring and evaluating student progress on IEP goals.

References

Bohannon, R. Direct and daily measurement procedures in the identification and treatment of reading behaviors of children in special education. Unpublished doctoral dissertation, University of Washington, 1975.

Deno, S. L., Mirkin, P. K., Lowry, L, & Kuehnle, K. Relationships among simple measures of spelling and performance on standardized achievement tests (Research Report No. 20). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities, 1980.

Federal Register. Department of Health, Education, and Welfare, Washington, D.C., 42(163), Tuesday, August 23, 1977, Part II.

Haring, N. G., & Berman, M. L. Evaluation and specification of instructional materials and management procedures through use of a pupil performance profile and a predictive model. Washington, D.C.: Bureau of Education for the Handicapped, 1973.

Harris, A. J., & Jacobson, M. D. Basic elementary reading vocabularies - The first R series. New York: MacMillan, 1972.

Jenkins, J. R., Deno, S. L., & Mirkin, P. K. Measuring pupil progress toward the least restrictive alternative. Learning Disability Quarterly, 1979, 2(4), 81-92.

Klein, S., Fenstermacher, G., & Aiken, M. C. The center's changing evaluation model. Evaluation Comment, 1971, 2(4), 9-12.

Liberty, K. A. Data decision rules (Working Paper #20). Eugene, Ore.: Regional Resources Center, University of Oregon, 1972.

Liberty, K. A. Decide for progress: Dynamic aims and data decisions. Seattle: Experimental Education Unit, Child Development and Mental Retardation Center, University of Washington, 1975.

Lindsley, O. R. Precision teaching in perspective: An interview with Ogden R. Lindsley. Teaching Exceptional Children, 1971, 3(3), 114-119.

Lovitt, T. C. Applied behavior analysis and learning disabilities: Part 1: Characteristics of ABA, general recommendations, and methodological limitations. Journal of Learning Disabilities, 1975, 8(7), 432-443.

Mirkin, P. K., & Deno, S. L. Formative evaluation in the classroom: An approach to improving instruction (Research Report No. 10). Minneapolis: University of Minnesota; Institute for Research on Learning Disabilities, 1979.

Scriven, M. The methodology of evaluation. In R. Tyler, R. Gagne, & M. Scriven (Eds.), Perspectives of curriculum evaluation (American Educational Research Association Monograph Series on Curriculum Evaluation). Chicago: Rand McNally, 1967.

Starlin, C. The use of daily direct recording as an aid in teaching oral reading. Unpublished doctoral dissertation, University of Oregon, 1970.

White, O. R. Evaluating the educational process (Working Paper). Seattle: University of Washington, Experimental Education Unit, 1974.

White, O. R., & Haring, N. G. Exceptional teaching: A multimedia training package. Columbus, Ohio: Charles E. Merrill, 1976.

White, O. R., & Liberty, K. A. Practical classroom measurement: Precision teaching (Working Paper). Seattle: University of Washington, Experimental Education Unit, 1974.

Table 1

Means and Standard Deviations of Student Rate Data for First Treatment Period

	Letter Sequences			Words		
	$\bar{X}$	$\bar{X}$ adj	SD	$\bar{X}$	$\bar{X}$ adj	SD
Grade Specific Test						
DMDB	31.6	30.9	11.4	3.1	3.1	1.6
DMTJ	32.8	32.2	10.9	3.2	3.4	1.7
WMTJ	24.6	25.8	10.2	2.5	2.4	1.0
Grand Master Test						
DMDB	24.0	22.9	10.9	6.0	5.5	3.5
DMTJ	23.7	22.5	7.4	5.2	5.1	2.3
WMTJ	17.3	19.4	9.2	4.3	4.9	3.2

Table 2

Daily vs. Weekly Measurement Contrast on Student Rate Data for First Treatment Period[a]

	Letter Sequences		Words	
	t	p	t	p
Grade Specific Test	3.71	.000	2.90	.005
Grand Master Test	3.17	.002	.70	.48

[a]Contrast was between students in daily measurement conditions (DMDB and DMTJ) during the first treatment period and students in weekly measurement (WMTJ) during the first treatment period.

Table 3

Means and Standard Deviations of Student Rate Data for Second Treatment Period

	Letter Sequences			Words		
	$\bar{X}$	$\bar{X}$ adj	SD	$\bar{X}$	$\bar{X}$ adj	SD
Grade Specific Test						
DMDB	32.6	36.1	11.0	4.0	4.2	1.7
DMTJ	37.8	34.5	10.5	3.6	3.6	1.8
WMTJ	31.5	30.7	12.1	3.4	3.2	1.5
Grand Master Test						
DMDB	22.0	26.1	9.6	5.3	6.2	3.1
DMTJ	29.0	24.5	9.7	7.6	6.6	3.5
WMTJ	23.0	23.0	9.0	4.8	4.9	2.5

Table 4

Daily vs. Weekly Measurement Contrast on Student Rate Data for Second Treatment Period[a]

	Letter Sequences		Words	
	t	p	t	p
Grade Specific Test	4.00	.000	2.85	.006
Grand Master Test	2.42	.02	2.97	.004

[a]Contrast was between students in daily measurement conditions during the second treatment period (DMDB and DMTJ) and students in weekly measurement during the second treatment period (WMTJ).

Table 5

Means and Standard Deviations of Student Rate Data on Grand Master Test Across Two Treatment Periods

Group	Letter Sequences $\bar{X}$	Letter Sequences $\bar{X}$ adj	Letter Sequences SD	Words $\bar{X}$	Words $\bar{X}$ adj	Words SD
Daily Measurement[a] Only	30.5	27.0	9.5	7.6	6.5	3.3
Daily and Weekly[b] Measurement	21.8	23.4	8.6	5.1	5.6	3.0

[a]This group included students in daily measurement during both treatment periods (DMDB-DMTJ or DMTJ-DMDB).

[b]This group included students in daily measurement during one treatment period and weekly measurement during the other treatment period (DMDB-WMTJ, WMTJ-DMDB, DMTJ-WMTJ, WMTJ-DMTJ).

Table 6

Daily Measurement Only vs. Daily and Weekly Measurement Contrast on Student Rate Data Across Two Treatment Periods

	Letter Sequences		Words	
	t	p	t	p
Grand Master List	3.10	.003	1.41	.16

Table 7

Data-Based Decision Rules vs. Teacher Judgment Contrast on Student Rate Data For First and Second Treatment Periods[a]

	Letter Sequences		Words	
	t	p	t	p
Grade Specific Test				
First treatment period	.69	.49	.92	.36
Second treatment period	1.23	.22	1.82	.07
Grand Master Test				
First treatment period	.33	.74	.72	.48
Second treatment period	1.41	.16	.59	.56

[a]Contrast was between students whose teachers used data-based decision rules (DMDB) and students whose teachers used their own judgment (DMTJ and WMTJ) for deciding upon program changes.

Table 8

Means and Standard Deviations of Student Rate Data on Grade Specific Test for Each Treatment in Each Combination

Combinations/Treatments	Letter Sequences		Words	
	$\bar{X}$	SD	$\bar{X}$	SD
DMDB/DMTJ[a]				
DMDB	38.8	9.6	3.5	1.9
DMTJ	35.4	10.1	3.3	1.8
DMDB/WMTJ[b]				
DMDB	27.2	9.9	3.6	1.6
WMTJ	23.9	9.8	2.9	1.5
DMTJ/WMTJ[c]				
DMTJ	35.2	11.8	3.3	1.5
WMTJ	32.4	11.9	3.1	1.4

[a]This combination included students in both DMDB-DMTJ and DMTJ-DMDB.

[b]This combination included students in both DMDB-WMTJ and WMTJ-DMDB.

[c]This combination included students in both DMTJ-WMTJ and WMTJ-DMTJ.

Table 9

Comparisons of Treatments Within Treatment Combinations

for Student Rate Data on Grade Specific Test

Contrast[a]	Letter Sequences		Words	
	t	p	t	p
DMDB vs. DMTJ	2.05	.05	.74	.46
DMDB vs. WMTJ	3.43	.002	3.10	.004
DMTJ vs. WMTJ	1.77	.09	.86	.40

[a]The df for the three contrasts were 24, 29, and 24, respectively, down the table.

APPENDIX A

Grade Specific and Grand Master Tests

Grade Specific Test
1st treatment period

PREPRIMER

1. he
2. ball
3. dog
4. play
5. daddy
6. see
7. to
8. no
9. blue
10. look
11. work
12. can
13. here
14. who
15. will
16. is
17. go
18. do
19. we
20. a
21. big
22. something
23. my
24. look
25. come
26. make
27. call
28. not
29. help
30. in
31. red
32. stop
33. me
34. have
35. said
36. little
37. at
38. green
39. did
40. funny
41. are
42. mother
43. up
44. you
45. ride
46. with
47. and
48. get
49. it
50. for

Grade Specific Test
1st treatment period

A-2

PRIMER

1. like
2. all
3. cake
4. some
5. know
6. new
7. from
8. yes
9. thank
10. put
11. jump
12. your
13. one
14. too
15. of
16. pet
17. man
18. him
19. about
20. bike
21. went
22. paint
23. his
24. book
25. soon
26. show
27. around
28. on
29. ask
30. but
31. into
32. so
33. tree
34. house
35. home
36. fast
37. now
38. has
39. yellow
40. two
41. take
42. eat
43. sit
44. away
45. say
46. saw
47. she
48. goat
49. train
50. boat

Grade Specific Test
1st treatment period

A-3
Grade 1

1. long	26. last	51. from	76. balloon
2. box	27. if	52. many	77. over
3. turtle	28. hold	53. began	78. story
4. penny	29. were	54. came	79. talk
5. stopped	30. sat	55. animal	80. cage
6. happy	31. gave	56. think	81. road
7. as	32. live	57. more	82. dark
8. wagon	33. color	58. please	83. fly
9. when	34. cow	59. airplane	84. letter
10. night	35. lost	60. does	85. name
11. school	36. which	61. wet	86. pig
12. wish	37. cry	62. try	87. didn't
13. just	38. fat	63. miss	88. feet
14. took	39. town	64. sang	89. cat
15. horse	40. was	65. bag	90. right
16. four	41. pocket	66. made	91. where
17. head	42. should	67. sing	92. farm
18. white	43. back	68. find	93. rain
19. under	44. off	69. still	94. them
20. tomorrow	45. other	70. there	95. those
21. peanut	46. own	71. walk	96. hurry
22. won't	47. after	72. am	97. kind
23. than	48. ice	73. I'll	98. shoe
24. coat	49. fall	74. ready	99. hop
25. laugh	50. hat	75. old	100. maybe

Grade Specific Test
1st treatment period

A-4

Grade 2

1. family	26. hung	51. people	76. third
2. yard	27. slow	52. lick	77. lunch
3. open	28. that's	53. life	78. point
4. rest	29. giant	54. second	79. anything
5. wood	30. honey	55. knock	80. butter
6. close	31. woman	56. lake	81. neck
7. once	32. shovel	57. air	82. winter
8. quick	33. bowl	58. wonder	83. wolf
9. bottom	34. policeman	59. ever	84. bell
10. kept	35. anyone	60. win	85. stand
11. arm	36. listen	61. I'm	86. king
12. feed	37. song	62. tent	87. pop
13. stick	38. fix	63. glass	88. wait
14. pass	39. roof	64. dirty	89. beside
15. blueberry	40. breakfast	65. trouble	90. wide
16. quack	41. idea	66. pretty	91. park
17. save	42. bought	67. done	92. large
18. through	43. return	68. stream	93. fierce
19. forest	44. until	69. year	94. safe
20. set	45. hot	70. week	95. nice
21. cherry	46. tower	71. spot	96. also
22. remember	47. goose	72. end	97. pick
23. whistle	48. buy	73. top	98. blew
24. gray	49. number	74. wrong	99. turn
25. apartment	50. easy	75. son	100. swim

Grade Specific Test
1st treatment period

A-5
Grade 3

1. nail	26. princess	51. happenings	76. danger
2. attention	27. moment	52. nibble	77. written
3. camera	28. speech	53. noon	78. measure
4. aboard	29. final	54. awake	79. enormous
5. straighten	30. buffalo	55. plain	80. dine
6. thump	31. shake	56. scream	81. shy
7. treasure	32. diamond	57. balcony	82. tender
8. history	33. shade	58. perhaps	83. tube
9. father	34. mud	59. pretend	84. shoulder
10. tale	35. cool	60. eager	85. upside
11. knee	36. crowd	61. sandwich	86. bend
12. sky	37. tremble	62. dim	87. silent
13. hotel	38. report	63. during	88. disappear
14. silence	39. bean	64. refuse	89. salt
15. welcome	40. rooster	65. marry	90. business
16. skyscraper	41. fifth	66. band	91. whale
17. tongue	42. leap	67. flutter	92. appear
18. lad	43. sugar	68. drum	93. purple
19. surround	44. chop	69. age	94. colt
20. thud	45. war	70. machine	95. upward
21. taste	46. cottage	71. parrot	96. helicopter
22. useful	47. less	72. spear	97. carrot
23. jail	48. gay	73. overhead	98. fault
24. dug	49. lamp	74. sill	99. cape
25. weight	50. sigh	75. wrap	100. mama

Grade Specific Test
1st treatment period

A-6

Grade 4

1.	forever	26.	flashlight	51.	tramp	76.	happening
2.	justice	27.	breeze	52.	ranger	77.	bucket
3.	peculiar	28.	officer	53.	crate	78.	action
4.	rip	29.	provide	54.	object	79.	slim
5.	arrest	30.	county	55.	ordinary	80.	mild
6.	cot	31.	freeze	56.	nowadays	81.	scarce
7.	we'd	32.	pace	57.	putter	82.	tilt
8.	fist	33.	imitate	58.	section	83.	speechless
9.	banker	34.	memory	59.	mound	84.	aware
10.	remarkable	35.	wedding	60.	spun	85.	marvel
11.	peak	36.	ha	61.	slung	86.	mane
12.	gradual	37.	rifle	62.	zebra	87.	creature
13.	buggy	38.	prefix	63.	unload	88.	tinkle
14.	drawn	39.	setting	64.	schoolroom	89.	stroke
15.	rapid	40.	grumble	65.	knot	90.	chart
16.	umpire	41.	velvet	66.	murmur	91.	curve
17.	preparation	42.	web	67.	announce	92.	link
18.	fallen	43.	mistress	68.	intend	93.	cinnamon
19.	roam	44.	patience	69.	sank	94.	terribly
20.	duty	45.	slender	70.	conversation	95.	seventy
21.	nation	46.	command	71.	leash	96.	event
22.	limit	47.	forehead	72.	persuade	97.	gentlemen
23.	hush	48.	port	73.	terror	98.	entire
24.	mug	49.	modern	74.	material	99.	destroy
25.	sword	50.	flip	75.	deck	100.	wreck

Grade Specific Test
1st treatment period

A-7
Grade 5

1.	porter	26.	magician	51.	theory	76.	huckleberry
2.	brim	27.	monster	52.	legal	77.	butte
3.	require	28.	wasp	53.	fake	78.	title
4.	otter	29.	talent	54.	plump	79.	sock
5.	gleam	30.	howdy	55.	starboard	80.	armor
6.	kerchief	31.	explosion	56.	stray	81.	foster
7.	goal	32.	interview	57.	gasoline	82.	athletic
8.	rudder	33.	fare	58.	oriole	83.	dandy
9.	ferry	34.	double	59.	capsule	84.	mow
10.	yonder	35.	lily	60.	minister	85.	resource
11.	saint	36.	toast	61.	disappointment	86.	oat
12.	prospect	37.	snip	62.	outlaw	87.	survive
13.	license	38.	rate	63.	preserve	88.	spit
14.	hydrogen	39.	military	64.	gale	89.	attitude
15.	cluck	40.	pea	65.	sole	90.	border
16.	lance	41.	spire	66.	disguise	91.	artist
17.	wretched	42.	elder	67.	dismiss	92.	leopard
18.	gravity	43.	super	68.	agent	93.	spur
19.	shrink	44.	outsmart	69.	motionless	94.	performance
20.	compartment	45.	moose	70.	lurk	95.	propose
21.	volunteer	46.	flung	71.	unfriendly	96.	commotion
22.	junk	47.	vacuum	72.	distress	97.	goggles
23.	neigh	48.	stricken	73.	floe	98.	project
24.	blur	49.	shilling	74.	linger	99.	legislature
25.	straddle	50.	loudspeaker	75.	ford	100.	pageant

Grade Specific Test
1st treatment period

A-8

Grade 6

1.	exceeding	26.	curtsy	51.	overnight	76.	pike
2.	crib	27.	goad	52.	gait	77.	response
3.	musn't	28.	burglar	53.	butt	78.	wisp
4.	nip	29.	device	54.	tone	79.	tribute
5.	mower	30.	accomplishment	55.	plateau	80.	massacre
6.	poodle	31.	sedan	56.	ember	81.	outstanding
7.	render	32.	manual	57.	quail	82.	secrecy
8.	possess	33.	spout	58.	dingy	83.	consequence
9.	mercury	34.	vain	59.	misfortune	84.	extraordinary
10.	slumber	35.	rehearse	60.	joint	85.	undergrowth
11.	sentry	36.	tassel	61.	homework	86.	perspiration
12.	wig	37.	sketch	62.	midget	87.	measles
13.	generation	38.	assume	63.	stout	88.	thatch
14.	trait	39.	layer	64.	estimate	89.	positive
15.	relationship	40.	laundry	65.	brow	90.	impact
16.	inquiring	41.	eyelid	66.	mystify	91.	supermarket
17.	midwest	42.	infection	67.	pluck	92.	nick
18.	ancestry	43.	papyrus	68.	countryside	93.	routine
19.	illness	44.	pursuit	69.	federal	94.	spice
20.	reduce	45.	appetite	70.	obtain	95.	woolen
21.	convenience	46.	countdown	71.	hub	96.	blister
22.	irregular	47.	unravel	72.	maintain	97.	majestic
23.	parlor	48.	erect	73.	value	98.	source
24.	superior	49.	awoke	74.	gleeful	99.	collide
25.	liable	50.	privilege	75.	puma	100.	slush

Grade Specific Test
2nd treatment period

A-9
Grade 1

1. right	26. truck	51. time	76. try
2. where	27. told	52. found	77. bird
3. farm	28. brown	53. water	78. happy
4. those	29. came	54. turtle	79. barn
5. tomorrow	30. kitten	55. sang	80. window
6. hurry	31. oh	56. fat	81. hear
7. kind	32. gone	57. find	82. by
8. shoe	33. am	58. money	83. back
9. hop	34. pan	59. men	84. laugh
10. horse	35. took	60. head	85. guess
11. own	36. far	61. just	86. when
12. more	37. could	62. talk	87. party
13. maybe	38. pig	63. if	88. fox
14. black	39. five	64. cut	89. bark
15. pony	40. fly	65. be	90. sound
16. sat	41. hat	66. morning	91. there
17. any	42. box	67. good-by	92. these
18. thing	43. name	68. cold	93. story
19. I'll	44. met	69. them	94. cat
20. story	45. can't	70. street	95. catch
21. our	46. cry	71. again	96. as
22. think	47. won't	72. before	97. got
23. first	48. don't	73. were	98. food
24. rain	49. still	74. race	99. drop
25. prize	50. grass	75. wagon	100. does

Grade Specific Test
2nd treatment period

A-10 Grade 2

1. cowboy	26. tower	51. tiger	76. cream
2. rope	27. owl	52. seven	77. finish
3. through	28. roar	53. hall	78. careful
4. sand	29. ear	54. city	79. woman
5. pay	30. knew	55. above	80. stood
6. trot	31. riding	56. oven	81. past
7. spring	32. lamb	57. instead	82. open
8. field	33. late	58. visit	83. jay
9. trap	34. bad	59. bank	84. policeman
10. sell	35. we'll	60. silk	85. magic
11. follow	36. leave	61. empty	86. early
12. that's	37. pretty	62. side	87. write
13. choose	38. twice	63. fairy	88. inside
14. well	39. puppy	64. toward	89. beside
15. clean	40. son	65. basket	90. seed
16. sometime	41. dollar	66. dream	91. sudden
17. oil	42. blueberry	67. tractor	92. star
18. taken	43. queer	68. nice	93. cart
19. parade	44. bounce	69. joke	94. pink
20. lick	45. hole	70. plan	95. merry
21. trouble	46. breakfast	71. library	96. tie
22. year	47. bone	72. trunk	97. hide
23. meet	48. tip	73. scarecrow	98. across
24. pull	49. float	74. door	99. forgot
25. wait	50. save	75. almost	100. honey

Grade Specific Test
2nd treatment period

A-11
Grade 3

1.	baseball	26.	steer	51.	thief	76.	tongue
2.	lesson	27.	wake	52.	stool	77.	forward
3.	forget	28.	chimney	53.	order	78.	view
4.	shell	29.	somebody	54.	mill	79.	tug
5.	shone	30.	grin	55.	mutter	80.	bride
6.	begun	31.	orange	56.	proper	81.	spare
7.	judge	32.	snowstorm	57.	porcupine	82.	chest
8.	shade	33.	drove	58.	market	83.	bow
9.	born	34.	draw	59.	probably	84.	balcony
10.	women	35.	fifty	60.	tune	85.	brush
11.	village	36.	take	61.	broken	86.	million
12.	broom	37.	plenty	62.	bonnet	87.	pound
13.	leap	38.	thread	63.	excitement	88.	explain
14.	adventure	39.	less	64.	bead	89.	cool
15.	willing	40.	shoulder	65.	coin	90.	patch
16.	hatch	41.	gift	66.	sixteen	91.	lamp
17.	safety	42.	whisker	67.	clap	92.	master
18.	mama	43.	invisible	68.	experiment	93.	everybody
19.	frog	44.	band	69.	swallow	94.	tribe
20.	invite	45.	solve	70.	form	95.	majesty
21.	granny	46.	shore	71.	sweep	96.	raise
22.	island	47.	serious	72.	allow	97.	final
23.	apron	48.	pale	73.	fact	98.	inventor
24.	interesting	49.	refuse	74.	attention	99.	smart
25.	waste	50.	age	75.	whole	100.	upstairs

Grade Specific Test
2nd treatment period

A-12

Grade 4

1.	scoop	26.	alike	51.	moss	76.	whom
2.	approach	27.	spun	52.	messenger	77.	wade
3.	sped	28.	flashlight	53.	hinge	78.	wisdom
4.	furious	29.	replace	54.	level	79.	wedding
5.	destroy	30.	pacific	55.	acorn	80.	temperature
6.	mom	31.	remark	56.	possession	81.	streak
7.	mop	32.	breeze	57.	snug	82.	volcano
8.	astronaut	33.	notch	58.	painter	83.	swoop
9.	flush	34.	sprinkle	59.	frank	84.	hoe
10.	mistress	35.	eaten	60.	grape	85.	skip
11.	bald	36.	hawk	61.	ha	86.	framework
12.	rough	37.	nightfall	62.	justice	87.	inning
13.	relief	38.	constant	63.	somewhere	88.	crab
14.	basement	39.	concern	64.	innocent	89.	pupil
15.	freight	40.	deck	65.	glider	90.	gardenia
16.	being	41.	pinch	66.	jug	91.	pad
17.	roam	42.	imagination	67.	quiver	92.	scientific
18.	especially	43.	praise	68.	baseman	93.	anchor
19.	map	44.	frozen	69.	balance	94.	bid
20.	giggle	45.	mound	70.	creek	95.	prayer
21.	dolphin	46.	doubt	71.	arose	96.	loop
22.	flier	47.	ridge	72.	cab	97.	claim
23.	eagerness	48.	flicker	73.	mutt	98.	oxen
24.	shawl	49.	anteater	74.	bunch	99.	ax
25.	flight	50.	nonsense	75.	crocodile	100.	snake

Grade Specific Test
2nd treatment period

A-13

Grade 5

1.	eventual	26.	bay	51.	hail	76.	liar
2.	charm	27.	heckon	52.	emperor	77.	triumphant
3.	apricot	28.	hurtle	53.	determine	78.	yarn
4.	slosh	29.	grease	54.	trumpet	79.	ignore
5.	evil	30.	bluff	55.	legal	80.	bound
6.	endurance	31.	subway	56.	caw	81.	faithful
7.	slice	32.	permanent	57.	harsh	82.	distress
8.	civil	33.	minister	58.	experimental	83.	tourist
9.	gasoline	34.	boulder	59.	lily	84.	community
10.	gust	35.	legislature	60.	scarlet	85.	rugged
11.	halt	36.	series	61.	raisin	86.	fir
12.	brilliant	37.	corral	62.	blunder	87.	coconut
13.	clam	38.	warehouse	63.	descendant	88.	rigid
14.	captive	39.	gap	64.	sandal	89.	false
15.	project	40.	teem	65.	relay	90.	deerskin
16.	twilight	41.	wring	66.	plea	91.	zero
17.	pulley	42.	keg	67.	snout	92.	unable
18.	sensible	43.	melody	68.	capsule	93.	comic
19.	corridor	44.	install	69.	slamp	94.	brisk
20.	dismiss	45.	mathematics	70.	manufacture	95.	scorch
21.	moor	46.	governor	71.	shrewd	96.	daze
22.	ain't	47.	public	72.	telegram	97.	proceed
23.	confident	48.	plaid	73.	injure	98.	pouch
24.	toast	49.	lurch	74.	midst	99.	ramp
25.	pierce	50.	brittle	75.	soak	100.	brand

Grade Specific Test
2nd treatment period

A-14

Grade 6

1. ally	26. empress	51. dame	76. transparent
2. brew	27. prevention	52. khaki	77. trait
3. gorgeous	28. ooze	53. groom	78. standard
4. cushion	29. persist	54. conservation	79. unique
5. disk	30. slain	55. occupation	80. variation
6. mince	31. duckling	56. rotten	81. spectator
7. blubber	32. existence	57. pedestal	82. scheme
8. coarse	33. desperation	58. galleon	83. baton
9. lichen	34. destruction	59. analyze	84. archaeologist
10. lifeguard	35. console	60. eliminate	85. quest
11. parch	36. acrobat	61. ape	86. forecast
12. reasonable	37. alter	62. maize	87. dungeon
13. shuck	38. assemble	63. blare	88. rung
14. slack	39. actress	64. conserve	89. heed
15. hatred	40. runway	65. massive	90. descend
16. fabulous	41. isolate	66. binoculars	91. fluid
17. curry	42. bravo	67. hearth	92. friction
18. regret	43. estimate	68. sip	93. evident
19. maintain	44. population	69. critical	94. construction
20. employee	45. continual	70. ample	95. definite
21. sap	46. ravine	71. fulfill	96. grave
22. departure	47. reel	72. indigo	97. gash
23. prophet	48. batch	73. reflection	98. fleece
24. particle	49. oppose	74. overnight	99. playmate
25. hurdle	50. expanse	75. portion	100. jackal

words in isolation

List 1 A-15

ceiling	occur	gleam	hostile	secretary
finger	trip	stew	clover	pillar
dense	have	resolve	different	command
cage	pound	library	talk	sport
built	between	minister	girl	fern
button	industry	goal	right	they
golden	kept	brim	feet	trout
porter	bike	yet	kindle	shield
twist	rose	concrete	discussion	tunnel
brother	brave	fringe	register	barn
green	dungeon	familiar	require	pump
grind	pot	sample	modern	queer

words in isolation

List 2

am	deliberate	tie	ancient	start
notion	pet	coat	everyone	gym
shallow	halt	fish	pickle	pony
eastern	ruin	collect	disturb	never
estimate	prepare	sympathetic	fall	suspect
connect	cage	journal	marvelous	hid
honest	brain	at	oppose	injure
beyond	lamb	brief	robin	victory
grab	mirror	much	blew	interesting
kitty	farewell	chase	except	cost
hatred	condemn	under	solution	garbage
deep	fork	deal	value	automatic

words in isolation — List 3

forgotten	blue	what	got	shout
constant	had	an	saint	snow
buy	nice	person	saddle	front
man	ferry	rid	explorer	avenue
ourselves	pistol	show	chosen	pardon
bishop	central	space	robber	sway
plump	aim	scare	humble	beef
cluck	cold	summit	fan	been
lance	slam	direction	can	license
mysterious	prophecy	underground	rock	yonder
airplane	hurt	snare	mattress	stew
penetrate	scramble	wretched	fill	response

APPENDIX B

Examples of Daily Tests

Preprimer - Day 1

1.	my	3
2.	daddy	6
3.	who	4
4.	green	6
5.	work	5
6.	up	3
7.	it	3
8.	little	7
9.	to	3
10.	red	4
11.	blue	5
12.	come	5
13.	not	4
14.	a	2
15.	here	5
16.	he	3
17.	see	4
18.	you	4
19.	said	5
20.	look	5
21.	stop	5
22.	call	5
23.	no	3
24.	at	3
25.	this	5

Primer - Day 1

1.	all	4
2.	jump	5
3.	goat	5
4.	around	7
5.	so	3
6.	man	4
7.	two	4
8.	show	5
9.	fast	5
10.	ash	4
11.	house	6
12.	yellow	7
13.	bat	4
14.	went	5
15.	like	5
16.	new	4
17.	has	4
18.	paint	6
19.	thank	6
20.	around	7
21.	train	6
22.	say	4
23.	cake	5
24.	home	5
25.	some	5

Grade 1 - Day 1

1.	than	5
2.	stopped	8
3.	took	5
4.	kind	5
5.	school	7
6.	back	5
7.	rain	5
8.	fall	5
9.	when	5
10.	as	3
11.	wet	4
12.	gone	5
13.	more	5
14.	am	3
15.	cat	4
16.	laugh	6
17.	road	5
18.	came	5
19.	animal	7
20.	does	5
21.	still	6
22.	hold	5
23.	tomorrow	9
24.	just	5
25.	many	5

Grade 2 - Day 1

1.	blew	5
2.	policeman	10
3.	roof	5
4.	honey	6
5.	cherry	7
6.	family	7
7.	yard	5
8.	close	6
9.	fierce	7
10.	ever	5
11.	arm	4
12.	fix	4
13.	neck	5
14.	remember	9
15.	pop	4
16.	bottom	7
17.	son	4
18.	slow	5
19.	win	4
20.	number	7
21.	large	6
22.	third	6
23.	hot	4
24.	open	5
25.	anyone	7

Grade 3 - Day 1

1.	shy	4
2.	mama	5
3.	age	4
4.	tube	5
5.	less	5
6.	marry	6
7.	written	8
8.	dug	4
9.	enormous	9
10.	upside	7
11.	hotel	6
12.	shoulder	9
13.	balcony	8
14.	sugar	6
15.	leap	5
16.	tale	5
17.	carrot	7
18.	thump	6
19.	disappear	10
20.	cool	5
21.	business	9
22.	colt	5
23.	jail	5
24.	father	7
25.	sandwich	9

Grade 4 - Day 1

1.	slim	5
2.	banker	7
3.	deck	5
4.	forever	8
5.	link	5
6.	object	7
7.	forehead	9
8.	destroy	8
9.	hush	5
10.	mound	6
11.	peculiar	9
12.	creature	9
13.	provide	8
14.	sank	5
15.	officer	8
16.	justice	8
17.	bucket	7
18.	tramp	6
19.	zebra	6
20.	tilt	5
21.	fist	5
22.	rip	4
23.	leash	6
24.	conversation	13
25.	port	5

Grade 5 - Day 1

1.	performance	12
2.	border	7
3.	wretched	9
4.	floe	5
5.	fare	5
6.	distress	9
7.	explosion	10
8.	junk	5
9.	goggles	8
10.	leopard	8
11.	yonder	7
12.	ford	5
13.	sole	5
14.	gale	5
15.	agent	6
16.	ferry	6
17.	armor	6
18.	lily	5
19.	capsule	8
20.	outsmart	9
21.	straddle	9
22.	preserve	9
23.	mow	4
24.	propose	8
25.	elder	6

Grade 6 - Day 1

1.	laundry	8
2.	infection	10
3.	tribute	8
4.	value	6
5.	misfortune	11
6.	liable	7
7.	hub	4
8.	maintain	9
9.	accomplishment	15
10.	rehearse	9
11.	positive	9
12.	ancestry	9
13.	plateau	8
14.	curtsy	7
15.	erect	6
16.	possess	8
17.	midget	7
18.	render	7
19.	brow	5
20.	consequence	12
21.	crib	5
22.	pursuit	8
23.	massacre	9
24.	sedan	6
25.	outstanding	12

APPENDIX C

Scoring Procedures for Correct Letter Sequences

SCORING PROCEDURES

Two methods for scoring the student's spelling are available. For purposes of graphing and monitoring student progress, there needs to be a certain level of proficiency, above or below which performance is difficult to assess. The choice of either method should be based on the student's current skill level and on Tables 4 and 5 to determine what is most appropriate. For students whose spelling/writing performance includes the use of whole words and can accomodate a rate of at least 5-10 words per minute, the number of words spelled correctly and incorrectly should be counted. If the student is spelling at a level less than this, there aren't enough responses to provide a sensitive measure of his/her progress in spelling and a different method should be used. For these students, who are low level enough to preclude the use of whole words or accommodate a rate of less than five words spelled in one minute, letters in correct sequence should be counted.

Directions for Scoring Letters in Correct Sequence

This method focuses on every pair of letters which appear correctly together. Spell: Direct

1) Every word must have a beginning letter, which implicitly means that no other letter appears prior to the first. That is, there is a blank space at the beginning of the word. If the word begins with a "D," place a caret (inverted v) over the blank space and the "D."

2) If the next letter to follow "D" is an "I," the two letters DI are in the correct sequence; place a caret so that it joins the "D" and the "I."

___ D^I

3) If the next letter is an "R," again the two letters IR are in the correct sequence. Repeat Step 2 for letters IR.

___ D I^R

4) This process is repeated for each pair of letters until the entire word is scored. As in the blank space implicit in beginning each word, the word must end in the correct letter being followed by a blank space.

5) The following is the correct way to score the entire word- direct

___^D^I^R^E^C^T^___

6) As can be seen, with the correct spelling there are 7 letters in correct sequence. For any word that is spelled correctly, there will be one more caret (letter in correct sequence) than there are letters in the word. As in this example, the word - direct - has six letters in it. Therefore, there will be 7 letters in correct sequence if the word is spelled correctly in its entirety. This is because of the point given for beginning and ending the word in a blank.

Further example: A misspelling of the word "Direkt"

___^D^I^R^E K T^___

There are 5 letters in correct sequence and 2 in incorrect sequence.

APPENDIX D

Spelling Intervention Checklist

Study VIII

Week of _____

Spelling Interventions

	MONDAY			TUESDAY			WEDNESDAY			THURSDAY			FRIDAY		
	Indi-vidual	Competitive	Cooperative	Indi-vidual	Competitive	Cooperative	Indi-vidual	Competitive	Cooperative	Indi-vidual	Competitive	Cooperative	Indi-vidual	Competitive	Cooperative
1) Dictation															
Oral Response															
Whole words															
Phonic components															
Syllables															
Written Response															
Whole words															
Phonic components															
Syllables															
Sentences															
2) Copying															
Whole words															
Sentences															
3) Verbal Rehearsal i.e. language master															
4) Syllabication															
5) Underlining phonic elements															
6) Fill in missing words i.e. cloze															
7) Editing															
8) Tracing															
9) Multiple Choice															
10) Spelling Games															
11) Definitions															
12) Workbook															

APPENDIX E

Training Materials

Data Decision Rules

Overview:

In this condition, you will be asked to follow a few basic guidelines for determining when a student's progress is sufficient for you to continue using the same teaching methods or whether a new strategy might be necessary to keep performance from dropping below its current level or remaining at the same level. The procedures will involve determining baseline performance, establishing an aimline, plotting the student's performance on a daily spelling measure on a graph, and comparing it to the aimline.

Specific Procedures:

1) On the first day, a list of words will be dictated three times to each student. The median performance (the middle score after ranking the three scores in order) will be determined and plotted on the graph for day one. This will serve as the baseline.
2) Following administration of this list, an aimline will be plotted on the graph for a three-week period. This aimline indicates the required daily performance for each student if s/he is to achieve a 30% improvement rate in spelling performance each week during the three-week treatment period. Steps 1 and 2 will be completed by the experimenters.
3) You will then be given a list of 100 words to teach. For the next 15 days, the general procedure will involve teaching the words from this list to each of your students for 10 minutes every day, giving them a short spelling test immediately after

each instructional period, and plotting their performance on the graph. It is important to maintain a 10 minute period for instruction, a three minute measurement period, and approximately two minutes for correction.

4) The spelling measure consists of a list of 25 words which you will dictate to the student for a three minute period. You will have a different list for each day. In the dictation, each word should be presented as soon as the student has finished the previous word, with a maximum of 15 seconds allowed for each word. That is, if the student has not finished a word within 15 seconds, you should present the next word anyway.

5) The measure of spelling performance is the total number of letters in correct sequence and the total number of letters in incorrect sequence. If a student doesn't write a word, all letters in that word would be considered incorrect. Additions and omissions of letters are also counted as errors. Refer to your scoring sheet for additional information.

6) The total number of letters in correct sequence and the total number of letters in incorrect sequence should then be plotted on the graph. The test should be corrected each day in front of the student and should take approximately two minutes. It is important not to share the graph of the student's progress with the student.

7) Each day you should compare the student's performance on the daily spelling measure with his or her expected performance (aimline) using the following criteria:

a) If on three consecutive days, data points are plotted above the aimline, or on three consecutive days, data points are plotted below the aimline, a new aimline should be drawn on the graph as follows:

1. Find the median performance level obtained on those three days and mark this point on the previous day.
2. Draw a line parallel to the original aimline, extending to the end of the three-week period. This procedure should be used every time your student's performance is above or below the aimline for three days in a row. If a student performs above aimline for two days and at (or below) aimline on the third day or vice versa, no new aimline is needed.

b) If the student's performance falls below the aimline for three consecutive days, in addition to drawing a new aimline a different teaching strategy should also be implemented. However, there are no constraints on the type of instruction or particular materials adopted. If performance is above the line, no new teaching strategy should be implemented.

8) This adjustment of the aimline and consequent changing of teaching strategies when the student's performance is below what is expected should continue for the duration of the three weeks. It is important to maintain a ten minute instructional period, a three minute measurement period, and a two minute correction period at all times.

9) At the end of the treatment session, there will be a few days of transition when experimenters will give measures to each of the students. At this time, you will be assigned to the next condition and will be given a detailed explanation of the procedures to follow.

Teacher Judgment with Daily Measurement and Daily Graphing

Overview:

In this condition you will be asked to make the judgment of whether a student's progress is sufficient for you to continue using the same teaching methods or whether a new strategy might be necessary to keep performance from dropping below or maintaining its current level. The student's performance on a daily spelling measure will be plotted on a graph to help you make your decisions.

Specific Procedures:

1) On the first day, a list of words will be dictated three times to each student. The median performance (the middle score after ranking the three scores in order) will be determined and plotted on the graph for day one. This will serve as the baseline and will be completed by the experimenters.
2) You will then be given a list of 100 words to teach. For the next 15 days, the general procedure will involve teaching the words from this list to each of your students for ten minutes every day, giving them a short spelling test immediately after each instructional period, and plotting their performance on the graph. It is important to maintain a 10 minute instructional period, a three minute testing period, and approximately two minutes for correction.
3) The spelling measure consists of a list of words which you should dictate to the student for a three minute period. In the dictation, each word should be presented as soon as the

student has finished the previous word, with a maximum of 15 seconds allowed for each word. If the student does not finish a word within 15 seconds, you should present the next word anyway.

4) The measure of spelling performance is the total number of letters in correct sequence and the total number of letters in incorrect sequence. If a student doesn't write a word, all letters in that word would be considered incorrect. Additions and omission of letters are also counted as errors. Refer to your scoring sheet for additional information. The total number of letters in correct and incorrect sequence should then be plotted on the graph. The test should be corrected each day in front of the student, but it is important not to share the graph of the student's progress with the student.

5) Every day, you should compare the student's performance with your expectations for their progress and determine whether the instructional program is appropriate.

   If performance is less than you would like, a different teaching strategy should be implemented and continued until another strategy is deemed necessary.

   If performance is not below your expectations, nothing needs to be changed and you should continue with the same teaching strategies. However, there are no constraints on when to adopt different teaching strategies or on the particular methods used.

6) At the end of the treatment session, there will be a few days of transition when experimenters will give measures to each of the students. At this time, you will be assigned to the next condition and will be given a detailed explanation of the procedures to follow.

E-8

Teacher Judgment with Weekly Measurement

Overview:

In this condition, you will be asked to make the judgment of whether a student's progress is sufficient for you to continue using the same teaching methods or whether a new strategy might be necessary to keep performance from dropping below its current level or remaining at the same level. A spelling measure will be administered once a week and the results recorded in a performance book.

Specific Procedures:

1) On the first day, a list of words will be dictated three times to each student. The median performance (the middle score after ranking the three scores in order) will be determined and recorded for day one. This will serve as the baseline and will be completed by the experimenters.

2) You will then be given a list of 100 words to teach. For the next 15 days, the general procedure will involve teaching the words from this list to each of your students for fifteen minutes every day. You may teach as many words as you like and in any order you wish. Once each week you should administer a three minute spelling test to each student. The student's score on this test should be entered in your grade book.

3) You may choose the list of words to use for each week's spelling test. On the day the spelling test is given, you should teach for 10 minutes and use the other five minutes to test and correct the student's paper. Dictate the spelling list to the student for a three minute period. In the dictation, each word should

be presented as soon as the student has finished the previous word, with a maximum of 15 seconds allowed for each word. That is, if the student has not finished a word within 15 seconds, you should present the next word anyway.

4) The measure of spelling performance is the total number of letters in correct sequence and the total number of letters in incorrect sequence. If a student doesn't write a word, all the letters in that word would be considered incorrect. Additions and omissions of letters are also counted as errors. The test should be corrected in front of the student and should take approximately two minutes. No more than 15 minutes a day should be spent on instruction in spelling during the study.

5) At the end of each week you have an opportunity to decide whether the student's performance is such that a change in teaching strategy is necessary.

   If performance is less than you would like, you may make a change in teaching strategy or teaching materials. If performance is not below your expectations, nothing needs to be changed and you should continue with the same teaching strategies. However, there are no constraints on when to adopt different teaching strategies or on the particular methods used.

6) At the end of the 15 day treatment session there will be a few days of transition period when experimenters will give measures to each of the students. At this time you will be assigned to the next condition and given a detailed explanation of procedures to follow.

APPENDIX F

Additional Tables

Table 1

Means and Standard Deviations of

Teacher Rate Data for First Treatment Period and

Results of Daily vs. Weekly Measurement Contrast[a]

	Letter Sequences			Words		
	$\bar{X}$	$\bar{X}$ adj	SD	$\bar{X}$	$\bar{X}$ adj	SD
Grade Specific Test						
DMDB	32.2	31.6	5.7	3.2	3.1	0.9
DMTJ	34.6	33.3	10.8	3.3	3.6	1.5
WMTJ	24.8	27.0	9.8	2.4	2.3	0.8
Grand Master Test						
DMDB	23.7	23.2	8.0	6.2	5.8	2.1
DMTJ	23.8	22.4	6.2	5.4	5.3	1.6
MMTJ	17.7	19.7	8.0	4.2	4.7	2.5

	t	p	t	p
Grade Specific Test	2.53	.02	2.26	.04
Grand Master Test	1.97	.06	1.42	.17

[a]Contrast was between the scores for teachers in daily measurement (DMDB and DMTJ) during the first treatment period and the scores for teachers in weekly measurement (WMTJ) during the first treatment period.

Table 2

Means and Standard Deviations of

Student Proportion Data for First Treatment Period

and Results of Daily vs. Weekly Measurement Contrast[a]

	Letter Sequences			Words		
	$\bar{X}$	$\bar{X}$ adj	SD	$\bar{X}$	$\bar{X}$ adj	SD
Grade Specific Test						
DMDB	.72	.70	.15	.43	.44	.19
DMTJ	.73	.73	.12	.46	.48	.23
WMTJ	.67	.67	.09	.37	.35	.13
Grand Master Test						
DMDB	.60	.58	.12	.30	.27	.19
DMTJ	.57	.56	.09	.24	.23	.11
WMTJ	.51	.54	.12	.21	.25	.14

	t	p	t	p
Grade Specific Test	1.77	.08	2.71	.008
Grand Master Test	1.47	.15	.05	.96

[a]Contrast was between proportion scores of students in daily measurement conditions (DMDB and DMTJ) during the first treatment period and students in weekly measurement (WMTJ) during the first treatment period.

Table 3

Means and Standard Deviations of

Teacher Proportion Data for First Treatment Period

and Results of Daily vs. Weekly Measurement Contrast[a]

	Letter Sequences			Words		
	$\bar{X}$	$\bar{X}$ adj	SD	$\bar{X}$	$\bar{X}$ adj	SD
Grade Specific Test						
DMDB	.73	.71	.05	.44	.42	.11
DMTJ	.74	.74	.09	.47	.51	.18
WMTJ	.66	.68	.07	.36	.35	.07
Grand Master Test						
DMDB	.60	.58	.07	.29	.29	.10
DMTJ	.57	.55	.06	.24	.25	.07
WMTJ	.50	.55	.12	.20	.21	.10

	t	p	t	p
Grade Specific Test	1.87	.08	2.43	.03
Grand Master Test	.39	.70	.33	.74

[a]Contrast was between proportion scores for teachers in daily measurement conditions (DMDB and DMTJ) during the first treatment period and scores for teachers in weekly measurement (WMTJ) during the first treatment period.

Table 4

Means and Standard Deviations of

Teacher Rate Data for Second Treatment Period and

Results of Daily vs. Weekly Measurement Contrast[a]

	Letter Sequences			Words		
	$\bar{X}$	$\bar{X}$ adj	SD	$\bar{X}$	$\bar{X}$ adj	SD
Grade Specific Test						
DMDB	34.5	37.3	10.4	4.1	4.2	1.4
DMTJ	35.5	35.0	7.1	3.6	3.8	1.3
WMTJ	34.6	32.6	11.4	3.5	3.3	1.1
Grand Master Test						
DMDB	23.6	27.5	8.7	6.0	6.8	3.0
DMTJ	27.9	24.6	7.8	7.3	6.6	2.5
WMTJ	23.9	23.8	6.0	4.9	4.8	1.5

	t	p	t	p
Grade Specific Test	2.87	.009	2.32	.03
Grand Master Test	1.61	.12	2.34	.03

[a]Contrast was between the scores for teachers in daily measurement (DMDB and DMTJ) during the second treatment period and the scores for teachers in weekly measurement (WMTJ) during the second treatment period.

Table 5

Means and Standard Deviations of

Student Proportion Data for Second Treatment Period

and Results of Daily vs. Weekly Measurement Contrast[a]

	Letter Sequences			Words		
	$\bar{X}$	$\bar{X}$ adj	SD	$\bar{X}$	$\bar{X}$ adj	SD
Grade Specific Test						
DMDB	.79	.80	.13	.54	.54	.22
DMTJ	.77	.78	.10	.49	.49	.21
WMTJ	.74	.73	.10	.44	.44	.17
Grand Master Test						
DMDB	.56	.58	.10	.24	.27	.13
DMTJ	.59	.57	.14	.33	.29	.16
WMTJ	.59	.59	.10	.24	.25	.11

	Letter Sequences		Words	
	t	p	t	p
Grade Specific Test	3.12	.002	2.05	.05
Grand Master Test	.50	.62	1.32	.19

[a]Contrast was between proportion scores of students in daily measurement conditions (DMDB and DMTJ) during the second treatment period and students in weekly measurement (WMTJ) during the second treatment period.

Table 6

Means and Standard Deviations of

Teacher Proportion Data for Second Treatment Period

and Results of Daily vs. Weekly Measurement Contrast[a]

	Letter Sequences			Words		
	$\bar{X}$	$\bar{X}$ adj	SD	$\bar{X}$	$\bar{X}$ adj	SD
Grade Specific Test						
DMDB	.79	.80	.09	.55	.53	.16
DMTJ	.77	.78	.07	.48	.50	.15
WMTJ	.75	.74	.08	.45	.44	.13
Grand Master Test						
DMDB	.58	.59	.07	.27	.29	.10
DMTJ	.63	.62	.08	.30	.28	.11
WMTJ	.58	.58	.06	.24	.24	.06

	t	p	t	p
Grade Specific Test	2.52	.02	1.65	.11
Grand Master Test	.96	.35	1.37	.19

[a]Contrast was between proportion scores for teachers in daily measurement conditions (DMDB and DMTJ) during the second treatment period and scores for teachers in weekly measurement (WMTJ) during the second treatment period.

Table 7

Means and Standard Deviations of

Teacher Rate Data on Grand Master Test Across Two

Treatment Periods and Results of Daily Measurement Only

vs. Daily and Weekly Measurement Contrast

	Letter Sequences			Words		
Group	$\bar{X}$	$\bar{X}$ adj	SD	$\bar{X}$	$\bar{X}$ adj	SD
Daily Measurement[a] Only	30.1	27.2	6.2	7.7	6.9	2.4
Daily and Weekly Measurement[b]	22.6	24.2	6.9	5.2	5.7	2.2
Contrast		t	p	t	p	
Daily Measurement Only vs. Daily and Weekly Measurement		2.16	.04	1.44	.17	

[a]The group included teachers in daily measurement during both treatment periods (DMDB-DMTJ or DMTJ-DMDB).

[b]This group included teachers in daily measurement during one treatment period and weekly measurement during the other treatment period (DMDB-WMTJ, WMTJ-DMDB, DMTJ-WMTJ, WMTJ-DMTJ).

Table 8

Means and Standard Deviations of

Student Proportion Data on Grand Master Test Across Two

Treatment Periods and Results of Daily Measurement Only vs.

Daily and Weekly Measurement Contrast

	Letter Sequences			Words		
Group	$\bar{X}$	$\bar{X}$ adj	SD	$\bar{X}$	$\bar{X}$ adj	SD
Daily Measurement[a] Only	.60	.58	.13	.34	.30	.15
Daily and Weekly Measurement[b]	.57	.58	.10	.24	.26	.12

Contrast	t	p	t	p
Daily Measurement Only vs. Daily and Weekly Measurement	.23	.81	1.41	.16

[a]The group included students in daily measurement during both treatment periods (DMDB-DMTJ or DMTJ-DMDB).

[b]This group included students in daily measurement during one treatment period and weekly measurement during the other treatment period (DMDB-WMTJ, WMTJ-DMDB, DMTJ-WMTJ, WMTJ-DMTJ).

Table 9

Means and Standard Deviations of

Teacher Proportion Data on Grand Master Test Across Two

Treatment Periods and Results of Daily Measurement Only vs.

Daily and Weekly Measurement Contrast

	Letter Sequences			Words		
Group	X	X adj	SD	X	X adj	SD
Daily Measurement[a] Only	.65	.64	.06	.33	.30	.09
Daily and Weekly Measurement[b]	.57	.58	.06	.24	.25	.08

Contrast	t	p	t	p
Daily Measurement Only vs. Daily and Weekly Measurement	1.83	.08	1.06	.30

[a]This group included teachers in daily measurement during both treatment periods (DMDB-DMTJ or DMTJ-DMDB).

[b]This group included teachers in daily measurement during one treatment period and weekly measurement during the other treatment period (DMDB-WMTJ, WMTJ-DMDB, DMTJ-WMTJ, WMTJ-DMTJ).

Table 10

Data-Based Decision Rules vs. Teacher Judgment Contrast

on Student Proportion Data

For First and Second Treatment Periods[a]

	Letter Sequences		Words	
	t	p	t	p
Grade Specific Test				
First treatment period	1.06	.29	.98	.33
Second treatment period	1.09	.28	1.05	.29
Grand Master Test				
First treatment period	.96	.34	1.23	.22
Second treatment period	.61	.54	.72	.48

[a]Contrast was between students whose teachers used data-based decision rules (DMDB) and students whose teachers used their own judgment (DMTJ and WMTJ) for deciding upon program changes.

Table 11

Data-Based Decision Rules vs. Teacher Judgment Contrast

on Teacher Rate Data

For First and Second Treatment Periods[a]

	Letter Sequences		Words	
	t	p	t	p
Grade Specific Test				
First treatment period	.74	.47	1.03	.32
Second treatment period	1.58	.13	1.31	.21
Grand Master Test				
First treatment period	.47	.65	.84	.41
Second treatment period	1.65	.12	.22	.83

[a]Contrast was between teachers who used data-based decision rules (DMDB) and teachers who used their own judgment (DMTJ and WMTJ) for deciding upon program changes.

Table 12

Data-Based Decision Rules vs. Teacher Judgment Contrast

on Teacher Proportion Data

For First and Second Treatment Periods[a]

	Letter Sequences		Words	
	t	p	t	p
Grade Specific Test				
First treatment period	1.00	.33	1.67	.11
Second treatment period	1.06	.30	.57	.57
Grand Master Test				
First treatment period	.96	.35	1.10	.28
Second treatment period	1.05	.31	.20	.85

[a]Contrast was between teachers who used data-based decision rules (DMDB) and teachers who used their own judgment (DMTJ and WMTJ) for deciding upon program changes.

Table 13

Means and Standard Deviations of

Student Proportion Data on Grade Specific Test for Each

Treatment in Each Combination and Results of Treatment Contrasts

Combinations/Treatments	Letter Sequences		Words	
	$\bar{X}$	SD	$\bar{X}$	SD
DMDB/DMTJ[a]				
DMDB	.78	.11	.49	.21
DMTJ	.76	.13	.51	.24
DMDB/WMTJ[b]				
DMDB	.75	.14	.50	.22
WMTJ	.69	.11	.40	.16
DMTJ/WMTJ[c]				
DMTJ	.74	.09	.43	.17
WMTJ	.72	.10	.41	.16

Contrast[d]	Letter Sequences		Words	
	t	p	t	p
DMDB vs. DMTJ	.80	.43	.64	.53
DMDB vs. WMTJ	3.19	.003	3.26	.003
DMTJ vs. WMTJ	1.42	.17	.65	.52

[a]This combination included students in both DMDB-DMTJ and DMTJ-DMDB.

[b]This combination included students in both DMDB-WMTJ and WMTJ-DMDB.

[c]This combination included students in both DMTJ-WMTJ and WMTJ-DMTJ.

[d]The df for the three contrasts were 23, 29, and 23, respectively, down the table.

F-14

Table 14

Means and Standard Deviations of

Teacher Rate Data on Grade Specific Test for Each

Treatment in Each Combination and Results of Treatment Contrasts

Combinations/Treatments	Letter Sequences		Words	
	$\bar{X}$	SD	$\bar{X}$	SD
DMDB/DMTJ[a]				
DMDB	37.26	6.76	3.8	1.4
DMTJ	34.93	6.60	3.6	1.6
DMDB/WMTJ[b]				
DMDB	28.71	7.15	3.5	1.0
WMTJ	26.31	8.79	2.9	1.0
DMTJ/WMTJ[c]				
DMTJ	35.19	11.17	3.3	1.2
WMTJ	33.30	13.11	3.1	1.2

Contrast[d]	Letter Sequences		Words	
	t	p	t	p
DMDB vs. DMTJ	1.30	.23	.39	.71
DMDB vs. WMTJ	1.39	.21	1.51	.18
DMTJ vs. WMTJ	.99	.35	.69	.52

[a]This combination included teachers in both DMDB-DMTJ and DMTJ-DMDB.

[b]This combination included teachers in both DMDB-WMTJ and WMTJ-DMDB.

[c]This combination included teachers in both DMTJ-WMTJ and WMTJ-DMTJ.

[d]The df for the three contrasts were 7, 6, and 7, respectively, down the table.

Table 15

Means and Standard Deviations of

Teacher Proportion Data on Grade Specific Test for Each

Treatment in Each Combination and Results of Treatment Contrasts

Combinations/Treatments	Letter Sequences		Words	
	$\bar{X}$	SD	$\bar{X}$	SD
DMDB/DMTJ[a]				
DMDB	.77	.07	.48	.16
DMTJ	.76	.10	.51	.20
DMDB/WMTJ[b]				
DMDB	.75	.08	.50	.14
WMTJ	.70	.08	.41	.11
DMTJ/WMTJ[c]				
DMTJ	.74	.07	.44	.11
WMTJ	.71	.08	.41	.12

Contrast[a]	Letter Sequences		Words	
	t	p	t	p
DMDB vs. DMTJ	.52	.62	.68	.52
DMDB vs. WMTJ	1.72	.14	1.79	.12
DMTJ vs. WMTJ	1.24	.26	.79	.46

[a]This combination included teachers in both DMDB-DMTJ and DMTJ-DMDB.

[b]This combination included teachers in both DMDB-WMTJ and WMTJ-DMDB.

[c]This combination included teachers in both DMTJ-WMTJ and WMTJ-DMTJ.

[d]The df for the contrasts were 7, 6, and 7, respectively, down the table.

PUBLICATIONS

Institute for Research on Learning Disabilities
University of Minnesota

The Institute is not funded for the distribution of its publications. Publications may be obtained for $3.00 per document, a fee designed to cover printing and postage costs. Only checks and money orders payable to the University of Minnesota can be accepted. All orders must be prepaid.

Requests should be directed to: Editor, IRLD, 350 Elliott Hall; 75 East River Road, University of Minnesota, Minneapolis, MN 55455.

Ysseldyke, J. E. Assessing the learning disabled youngster: The state of the art (Research Report No. 1). November, 1977.

Ysseldyke, J. E., & Regan, R. R. Nondiscriminatory assessment and decision making (Monograph No. 7). February, 1979.

Foster, G., Algozzine, B., & Ysseldyke, J. Susceptibility to stereotypic bias (Research Report No. 3). March, 1979.

Algozzine, B. An analysis of the disturbingness and acceptability of behaviors as a function of diagnostic label (Research Report No. 4). March, 1979.

Algozzine, B., & McGraw, K. Diagnostic testing in mathematics: An extension of the PIAT? (Research Report No. 5). March, 1979.

Deno, S. L. A direct observation approach to measuring classroom behavior: Procedures and application (Research Report No. 6). April, 1979.

Ysseldyke, J. E., & Mirkin, P. K. Proceedings of the Minnesota roundtable conference on assessment of learning disabled children (Monograph No. 8). April, 1979.

Somwaru, J. P. A new approach to the assessment of learning disabilities (Monograph No. 9). April, 1979.

Algozzine, B., Forgnone, C., Mercer, C. D., & Trifiletti, J. J. Toward defining discrepancies for specific learning disabilities: An analysis and alternatives (Research Report No. 7). June, 1979.

Algozzine, B. The disturbing child: A validation report (Research Report No. 8). June, 1979.

Note: Monographs No. 1 - 6 and Research Report No. 2 are not available for distribution. These documents were part of the Institute's 1979-1980 continuation proposal, and/or are out of print.

Ysseldyke, J. E., Algozzine, B., Regan, R., & Potter, M. Technical adequacy of tests used by professionals in simulated decision making (Research Report No. 9). July, 1979.

Jenkins, J. R., Deno, S. L., & Mirkin, P. K. Measuring pupil progress toward the least restrictive environment (Monograph No. 10). August, 1979.

Mirkin, P. K., & Deno, S. L. Formative evaluation in the classroom: An approach to improving instruction (Research Report No. 10). August, 1979.

Thurlow, M. L., & Ysseldyke, J. E. Current assessment and decision-making practices in model programs for the learning disabled (Research Report No. 11). August, 1979.

Deno, S. L., Chiang, B., Tindal, G., & Blackburn, M. Experimental analysis of program components: An approach to research in CSDC's (Research Report No. 12). August, 1979.

Ysseldyke, J. E., Algozzine, B., Shinn, M., & McGue, M. Similarities and differences between underachievers and students labeled learning disabled: Identical twins with different mothers (Research Report No. 13). September, 1979.

Ysseldyke, J., & Algozzine, R. Perspectives on assessment of learning disabled students (Monograph No. 11). October, 1979.

Poland, S. F., Ysseldyke, J. E., Thurlow, M. L., & Mirkin, P. K. Current assessment and decision-making practices in school settings as reported by directors of special education (Research Report No. 14). November, 1979.

McGue, M., Shinn, M., & Ysseldyke, J. Validity of the Woodcock-Johnson psycho-educational battery with learning disabled students (Research Report No. 15). November, 1979.

Deno, S., Mirkin, P., & Shinn, M. Behavioral perspectives on the assessment of learning disabled children (Monograph No. 12). November, 1979.

Sutherland, J. H., Algozzine, B., Ysseldyke, J. E., & Young, S. What can I say after I say LD? (Research Report No. 16). December, 1979.

Deno, S. L., & Mirkin, P. K. Data-based IEP development: An approach to substantive compliance (Monograph No. 13). December, 1979.

Ysseldyke, J., Algozzine, B., Regan, R., & McGue, M. The influence of test scores and naturally-occurring pupil characteristics on psycho-educational decision making with children (Research Report No. 17). December, 1979.

Algozzine, B., & Ysseldyke, J. E. Decision makers' prediction of students' academic difficulties as a function of referral information (Research Report No. 18). December, 1979.

Ysseldyke, J. E., & Algozzine, B. Diagnostic classification decisions as a function of referral information (Research Report No. 19). January, 1980.

Deno, S. L., Mirkin, P. K., Chiang, B., & Lowry, L. Relationships among simple measures of reading and performance on standardized achievement tests (Research Report No. 20). January, 1980.

Deno, S. L., Mirkin, P. K., Lowry, L., & Kuehnle, K. Relationships among simple measures of spelling and performance on standardized achievement tests (Research Report No. 21). January, 1980.

Deno, S. L., Mirkin, P. K., & Marston, D. Relationships among simple measures of written expression and performance on standardized achievement tests (Research Report No. 22). January, 1980.

Mirkin, P. K., Deno, S. L., Tindal, G., & Kuehnle, K. Formative evaluation: Continued development of data utilization systems (Research Report No. 23). January, 1980.

Deno, S. L., Mirkin, P. K., Robinson, S., & Evans, P. Relationships among classroom observations of social adjustment and sociometric rating scales (Research Report No. 24). January, 1980.

Thurlow, M. L., & Ysseldyke, J. E. Factors influential on the psychoeducational decisions reached by teams of educators (Research Report No. 25). February, 1980.

Ysseldyke, J. E., & Algozzine, B. Diagnostic decision making in individuals susceptible to biasing information presented in the referral case folder (Research Report No. 26). March, 1980.

Thurlow, M. L., & Greener, J. W. Preliminary evidence on information considered useful in instructional planning (Research Report No. 27). March, 1980.

Ysseldyke, J. E., Regan, R. R., & Schwartz, S. Z. The use of technically adequate tests in psychoeducational decision making (Research Report No. 28). April, 1980.

Richey, L., Potter, M., & Ysseldyke, J. Teachers' expectations for the siblings of learning disabled and non-learning disabled students: A pilot study (Research Report No. 29). May, 1980.

Thurlow, M. L., & Ysseldyke, J. E. Instructional planning: Information collected by school psychologists vs. information considered useful by teachers (Research Report No. 30). June, 1980.

Algozzine, B., Webber, J., Campbell, M., Moore, S., & Gilliam, J. Classroom decision making as a function of diagnostic labels and perceived competence (Research Report No. 31). June, 1980.

"END OF DOCUMENT"